

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANTS: Hein, Mich B.
Hiatt, Andrew C.
Fitchen, John H.
- (ii) TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
- (iii) NUMBER OF SEQUENCES: 113
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SEED and BERRY LLP
(B) STREET: 6300 Columbia Center, 701 Fifth Avenue
(C) CITY: Seattle
(D) STATE: Washington
(E) COUNTRY: USA
(F) ZIP: 98104
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 09-JAN-1998
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Maki, David J.
(B) REGISTRATION NUMBER: 31,392
(C) REFERENCE/DOCKET NUMBER: 310098.401C1
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (206) 622-4900
(B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 137 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

05005318 010998

Gln	Glu	Asp	Glu	Arg	Ile	Val	Leu	Val	Asp	Asn	Lys	Cys	Lys	Cys	Ala
1				5					10					15	
Arg	Ile	Thr	Ser	Arg	Ile	Ile	Arg	Ser	Ser	Glu	Asp	Pro	Asn	Glu	Asp
			20					25					30		
Ile	Val	Glu	Arg	Asn	Ile	Arg	Ile	Ile	Val	Pro	Leu	Asn	Asn	Arg	Glu
		35					40					45			
Asn	Ile	Ser	Asp	Pro	Thr	Ser	Pro	Leu	Arg	Thr	Arg	Pro	Val	Tyr	His
	50					55					60				
Leu	Ser	Asp	Leu	Cys	Lys	Lys	Cys	Asp	Pro	Thr	Glu	Val	Glu	Leu	Asp
65					70					75					80
Asn	Gln	Ile	Val	Thr	Ala	Thr	Gln	Ser	Asn	Ile	Cys	Asp	Glu	Asp	Ser
				85					90					95	
Ala	Thr	Glu	Thr	Cys	Tyr	Thr	Tyr	Asp	Arg	Asn	Lys	Cys	Tyr	Thr	Ala
			100					105					110		
Val	Val	Pro	Leu	Val	Tyr	Gly	Gly	Glu	Thr	Lys	Met	Val	Glu	Thr	Ala
		115					120					125			
Leu	Thr	Pro	Asp	Ala	Cys	Tyr	Pro	Asp							
	130					135									

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

Gln	Asp	Glu	Asn	Glu	Arg	Ile	Val	Val	Asp	Asn	Lys	Cys	Lys	Cys	Ala
1				5					10					15	
Arg	Ile	Thr	Ser	Arg	Ile	Ile	Pro	Ser	Ala	Glu	Asp	Pro	Ser	Gln	Asp
			20					25					30		
Ile	Val	Glu	Arg	Asn	Val	Arg	Ile	Ile	Val	Pro	Leu	Asn	Ser	Arg	Glu
		35					40					45			
Asn	Ile	Ser	Asp	Pro	Thr	Ser	Pro	Met	Arg	Thr	Lys	Pro	Val	Tyr	His
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:3:

(A) LENGTH: 137 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

Asp	Asp	Glu	Ala	Thr	Ile	Leu	Ala	Asp	Asn	Lys	Cys	Met	Cys	Thr	Arg
1				5					10					15	
Val	Thr	Ser	Arg	Ile	Ile	Pro	Ser	Thr	Glu	Asp	Pro	Asn	Glu	Asp	Ile
			20					25					30		
Val	Glu	Arg	Asn	Ile	Arg	Ile	Val	Val	Pro	Leu	Asn	Asn	Arg	Glu	Asn
		35					40					45			
Ile	Ser	Asp	Pro	Thr	Ser	Pro	Leu	Arg	Arg	Asn	Pro	Val	Tyr	His	Leu
	50					55					60				
Ser	Asp	Val	Cys	Lys	Lys	Cys	Asp	Pro	Val	Glu	Val	Glu	Leu	Glu	Asp
65					70					75					80
Gln	Val	Val	Thr	Ala	Thr	Gln	Ser	Asn	Ile	Cys	Asn	Glu	Asp	Asp	Gly
				85					90					95	
Val	Pro	Glu	Thr	Cys	Tyr	Met	Tyr	Asp	Arg	Asn	Lys	Cys	Tyr	Thr	Thr
			100					105					110		
Met	Val	Pro	Leu	Arg	Tyr	His	Gly	Glu	Thr	Lys	Met	Val	Gln	Ala	Ala
		115					120					125			
Leu	Thr	Pro	Asp	Ser	Cys	Tyr	Pro	Asp							

130

135

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 136 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Glu Asp Glu Ser Thr Val Leu Val Asp Asn Lys Cys Gln Cys Val Arg
 1 5 10 15
 Ile Thr Ser Arg Ile Ile Arg Asp Pro Asp Asn Pro Ser Glu Asp Ile
 20 25 30
 Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Thr Arg Glu Asn
 35 40 45
 Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Glu Pro Lys Tyr Asn Leu
 50 55 60
 Ala Asn Leu Cys Lys Lys Cys Asp Pro Thr Glu Ile Glu Leu Asp Asn
 65 70 75 80
 Gln Val Phe Thr Ala Ser Gln Ser Asn Ile Cys Pro Asp Asp Asp Tyr
 85 90 95
 Ser Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Thr Leu
 100 105 110
 Val Pro Ile Thr His Arg Gly Val Thr Arg Met Val Lys Ala Thr Leu
 115 120 125
 Thr Pro Asp Ser Cys Tyr Pro Asp
 130 135

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

09005318-010590

Glu	Gln	Glu	Tyr	Ile	Leu	Ala	Asn	Asn	Lys	Cys	Lys	Cys	Val	Lys	Ile					
			5						10						15					
Ser	Ser	Arg	Phe	Val	Pro	Ser	Thr	Glu	Arg	Pro	Gly	Glu	Glu	Ile	Leu					
			20						25						30					
Glu	Arg	Asn	Ile	Gln	Ile	Thr	Ile	Pro	Thr	Ser	Ser	Arg	Met	Xaa	Ile					
			35						40						45					
Ser	Asp	Pro	Tyr	Ser	Pro	Leu	Arg	Thr	Gln	Pro	Val	Tyr	Asn	Leu	Trp					
			50						55						60					
Asp	Ile	Cys	Gln	Lys	Cys	Asp	Pro	Val	Gln	Leu	Glu	Ile	Gly	Gly	Ile					
65						70						75						80		
Pro	Val	Leu	Ala	Ser	Gln	Pro	Xaa	Xaa	Ser	Xaa	Pro	Asp	Asp	Glu	Cys					
			85						90						95					
Tyr	Thr	Thr	Glu	Val	Asn	Phe	Lys	Lys	Lys	Val	Pro	Leu	Thr	Pro	Asp					
			100						105						110					
Ser	Cys	Tyr	Glu	Tyr	Ser	Glu														
			115																	

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 128 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

Asn	Lys	Cys	Met	Cys	Thr	Arg	Val	Thr	Ala	Arg	Ile	Arg	Gly	Thr	Arg
1				5					10					15	
Glu	Asp	Pro	Asn	Glu	Asp	Ile	Val	Glu	Arg	Tyr	Ile	Arg	Ile	Asn	Val
			20					25					30		
Pro	Leu	Lys	Asn	Arg	Gly	Asn	Ile	Ser	Asp	Pro	Thr	Ser	Pro	Leu	Arg
		35					40					45			
Asn	Gln	Pro	Val	Tyr	His	Leu	Ser	Pro	Ser	Cys	Lys	Lys	Cys	Asp	Pro
	50					55					60				
Tyr	Glu	Asp	Gly	Val	Val	Thr	Ala	Thr	Glu	Thr	Asn	Ile	Cys	Tyr	Pro
65					70					75					80

Asp Gln Gly Val Pro Gln Ser Cys Arg Asp Tyr Cys Pro Glu Leu Asp
85 90 95

Arg Asn Lys Cys Tyr Thr Val Leu Val Pro Pro Gly Tyr Thr Gly Glu
100 105 110

Thr Lys Met Val Gln Asn Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
115 120 125

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 421 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..414

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAT CAG GAA GAT GAA CGT ATT GTT CTG GTT GAC AAC AAG TGC AAG TGT	48
Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys	
1 5 10 15	
GCT CGT ATT ACT TCT AGA ATC ATC CGT AGC TCA GAG GAC CCA AAT GAA	96
Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu	
20 25 30	
GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC GTC CCA CTG AAT AAC CGG	144
Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg	
35 40 45	
GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG CGC ACA CGC TTC GTA TAC	192
Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr	
50 55 60	
CAC CTG TCA GAT CTG TGT AAG AAG TGT GAT CCA ACA GAG GTA GAG CTG	240
His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu	
65 70 75 80	
GAC AAT CAG ATA GTC ACT GCG ACT CAA AGC AAC ATT TGC GAT GAG GAC	288
Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp	
85 90 95	
AGC GCT ACA GAA ACC TGC AGC ACC TAC GAT AGG AAC AAA TGC TAC ACG	336
Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr	
100 105 110	

09005318-040998

GCC GTG GTT CCG CTC GTG TAT GGT GGA GAG ACA AAA ATG GTG GAA ACT 384
 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
 115 120 125

GCC CTT ACG CCC GAT GCA TGC TAT CCG GAC TGAATTC 421
 Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
 130 135

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..213

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAT CAG AAG TGC AAG TGT GCT CGT ATT ACT TCT AGA ATC ATC CGT AGC 48
 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
 1 5 10 15

TCA GAG GAC CCA AAT GAA GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC 96
 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
 20 25 30

GTC CCA CTG AAT AAC CGG GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG 144
 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
 35 40 45

CGC ACA CGC TTC GTA TAC CAC CTG TCA GAT CTG TGT AAG AAG GAT GAG 192
 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Asp Glu
 50 55 60

GAC AGC GCT ACA GAA ACC TGC TG 215
 Asp Ser Ala Thr Glu Thr Cys
 65 70

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

00005318-010998

(2) INFORMATION FOR SEQ ID NO:10:

GATCAGAAGT GCAAGTGTGC TCGTATTACT T 31

(2) INFORMATION FOR SEQ ID NO:12:

- (A) LENGTH: 109 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCACCTACGA TAGGAACAAA TGCTACACGG CCGTG GTTCC GCTCGTGTAT GGTGGAGAGA 60

CAAAAATGGT GGAAACTGCC CTTACG CCG ATGCATGCTA CCCTGACTG 109

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 286 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..282

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAC	AAC	AAG	TGC	AAG	TGT	GCT	CGT	ATT	ACT	TCT	AGA	ATC	ATC	CGT	AGC	48
Asp	Asn	Lys	Cys	Lys	Cys	Ala	Arg	Ile	Thr	Ser	Arg	Ile	Ile	Arg	Ser	
15					20					25					30	

TCA GAG GAC CCA AAT GAA GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC 96
Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
35 40 45

GTC CCA CTG AAT AAC CGG GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG 144
Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
50 55 60

CGC ACA CGC TTC GTA TAC CAC CTG TCA GAT CTG TGT AAG AAG TGT GAT 192
Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Cys Asp
65 70 75

CCA ACA GAG GTA GAG CTG GAC AAT CAG ATA GTC ACT GCG ACT CAA AGC 240
Pro Thr Glu Val Glu Leu Asp Asn Gln Ile Val Thr Ala Thr Gln Ser
80 85 90

AAC ATT TGC GAT GAG GAC AGC GCT ACA GAA ACC TGC TAC TGA 282
Asn Ile Cys Asp Glu Asp Ser Ala Thr Glu Thr Cys Tyr *

105

286

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: CDS

- (B) LOCATION: 1..105

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GATCAGGAAG ATGAACGTAT TGTTCTGGTT GACAACAAGT GCAAGTGTGC TCGTATTACT	60
T	61

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCGATGACGA CGATAAGGCC CAAACGGAGA CCTGTACTGT TGCGCCTCGT GAACGGCAAAA 60
 ACTGCGGATT CCCGGAAGTA ACACCCTCTC AGTGCGCTAA TAAAGGCTGC TGTTTTGATG 120
 ACACGGTACG GGGCGTTCCG TGGTGCTTCT ACCCCAATAC AATTGACGTT CCGCCTGAAG 180
 AAGAGTGCGA GCCGTAAG 198

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys
 1 5 10 15
 Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu
 20 25 30
 Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg
 35 40 45
 Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr
 50 55 60
 His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu
 65 70 75 80
 Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp
 85 90 95
 Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr
 100 105 110
 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
 115 120 125

09005348-010998
 866070-8750060

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Asn Ile Cys Asp Glu Asp Ser Ala Thr Glu Thr Cys Tyr
85 90

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Leu Trp Thr
35

(D) TOPOLOGY: linear

Ala Arg Ile Thr Ser Arg
20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

Glu Phe
65

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 421 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTAGTCCTTC TACTTGATA ACAAGACCAA CTGTTGTTCA CGTTCACACG AGCATAATGA 60
 AGATCTTAGT AGGCATCGAG TCTCCTGGGT TTACTTCTAT ATCAGCTTGC ATTGTAGGCA 120
 TAGTAGCAGG GTGACTTATT GGCCCTCTTA TAGAGTCTAG GATGTTTCAGG CAACGCGTGT 180
 GCGAAGCATA TGGTGGACAG TCTAGACACA TTCTTCACAC TAGGTTGTCT CCATCTCGAC 240
 CTGTTAGTCT ATCAGTGACG CTGAGTTTCG TTGTAAACGC TACTCCTGTC GCGATGTCTT 300
 TGGACGTCGT GGATGCTATC CTTGTTTACG ATGTGCCGGC ACCAAGGCGA GCACATACCA 360
 CCTCTCTGTT TTTACCACCT TTGACGGGAA TGCGGGCTAC GTACGATAGG CCTGACTTAA 420
 G 421

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTAGTCTTCA CGTTCACACG AGCATAATGA AGATCTTAGT AGGCATCGAG TCTCCTGGGT 60
 TTACTTCTAT ATCAGCTTGC ATTGTAGGCA TAGTAGCAGG GTGACTTATT GGCCCTCTTA 120
 TAGAGTCTAG GATGTTTCAGG CAACGCGTGT GCGAAGCATA TGGTGGACAG TCTAGACACA 180
 TTCTTCCTAC TCCTGTCGCG ATGTCTTTGG ACGACTTAA 219

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTAGTAGGCA TCGAGTCTCC TGGGTTTACT TCTATATCAG CTTGCATTGT AGGCATAGTA 60

0005318.040998

GCAGGGTGAC TTATTGGCCC TCTTATAGAG TCTAGGATGT TCAGGCAACG CGTGTGCGAA 120

GCAATATGGTG GACAGTCTAG 140

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TCTTCACGTT CACACGAGCA TAATGAAGAT C 31

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 44 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ACACATTCTT CCTACTTCTC AGGCGATGTC TTTGGACGAC TTAA 44

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 117 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ACGTCGTGGA TGCTATCCTT GTTTACGATG TGCCGGCACC AAGGCGAGCA CATACCACCT 60

CTCTGTTTTT ACCACCTTTG ACGGGAATGC GGGCTACGTA CGATGGGACT GACTTAA 117

(2) INFORMATION FOR SEQ ID NO:33:

0005319.010999
B56070.2FE50060

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 282 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTGTTGTTCA CGTTCACACG AGCATAATGA AGATCTTAGT AGGCATCGAG TCTCCTGGGT	60
TTACTTCTAT ATCAGCTTGC ATTGTAGGCA TAGTAGCAGG GTGACTTATT GGCCCTCTTA	120
TAGAGTCTAG GATGTTTCAGG CAACGCGTGT GCGAAGCATA TGGTGGACAG TCTAGACACA	180
TTCTTCACAC TAGGTTGTCT CCATCTCGAC CTGTTAGTCT ATCAGTGACG CTGAGTTTCG	240
TTGTAAACGC TACTCCTGTC GCGATGTCTT TGGACGATGA CT	282

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 105 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATCTGTGTA AGAAGTGTGA TCCAACAGAG GTAGAGCTGG ACAATCAGAT AGTCACTGCG	60
ACTCAAAGCA ACATTTGCGA TGAGGACAGC GCTACACTTT GGACG	105

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 65 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

000534 01093
000070 01093

Table 1	
Parameter	Value
Initial population size	1000
Number of generations	1000
Selection method	Tournament
Crossover probability	0.8
Mutation probability	0.01
Elitism	Yes
Parallelization	Yes
Number of processors	4
Hardware	Intel Core i7-4790K, 16GB RAM
Software	Java 8, Eclipse IDE
Operating system	Windows 10
Test data source	UCI Machine Learning Repository
Number of features	18
Number of classes	2
Training set size	1500
Test set size	500
Accuracy	0.95
Precision	0.95
Recall	0.95
F1 score	0.95
Area under the curve (AUC)	0.95
Confusion matrix	See Table 2
ROC curve	See Figure 1
Decision boundary	See Figure 2
Feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High
Model precision	High
Model recall	High
Model F1 score	High
Model AUC	High
Model confusion matrix	See Table 2
Model ROC curve	See Figure 1
Model decision boundary	See Figure 2
Model feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High
Model precision	High
Model recall	High
Model F1 score	High
Model AUC	High
Model confusion matrix	See Table 2
Model ROC curve	See Figure 1
Model decision boundary	See Figure 2
Model feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High
Model precision	High
Model recall	High
Model F1 score	High
Model AUC	High
Model confusion matrix	See Table 2
Model ROC curve	See Figure 1
Model decision boundary	See Figure 2
Model feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High
Model precision	High
Model recall	High
Model F1 score	High
Model AUC	High
Model confusion matrix	See Table 2
Model ROC curve	See Figure 1
Model decision boundary	See Figure 2
Model feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High
Model precision	High
Model recall	High
Model F1 score	High
Model AUC	High
Model confusion matrix	See Table 2
Model ROC curve	See Figure 1
Model decision boundary	See Figure 2
Model feature importance	See Table 3
Model interpretability	High
Model robustness	High
Model generalization	High
Model scalability	High
Model maintainability	High
Model security	High
Model privacy	High
Model transparency	High
Model accountability	High
Model explainability	High
Model fairness	High
Model bias	Low
Model variance	Low
Model overfitting	Low
Model underfitting	Low
Model instability	Low
Model sensitivity	Low
Model specificity	High

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ACTTCGCTAC	TGCTGCTATT	CCGGGTTTGC	CTCTGGACAT	GACAACGCGG	AGCACTTGCC	60
GTTTTGACGC	CTAAGGGCCT	TCATTGTGGG	AGAGTCACGC	GATTATTTCC	GACGACAAAA	120
CTACTGTGCC	ATGCCCCGCA	AGGCACCACG	AAGATGGGGT	TATGTTAACT	GCAAGGCGGA	180
CTTCTTCTCA	CGCTCGGCAT	TCTTAA				206

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys
1 5 10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Glu Asn Leu Tyr Phe Gln Ser
1 5

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Lys Ala His Lys Val Asp Met Val Gln Tyr Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Val Gln Tyr Thr
1

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Glu Lys Ala Val Ala Asp

0200313 010993
056010 010993

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Lys Asp Glu Leu
1

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Ala Ile Gln Asp Pro Arg Leu Phe Ala Glu Glu Lys Ala Val Ala Asp
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GATCAGGAAG ATGAACGTAT TGTTCTGGTT GACAACAAGT GCAAGTGTGC TCGTATTACT 60
T 61

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

0005318 010000

CTAGAAGTAA TACGAGCACA CTTGCACTTG TTGTCAACCA GAACAATACG TTCATCTTCC	60
T	61

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GATCAGAAGT GCAAGTGTGC TCGTATTACT T 31

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

CTAGAAGTAA TACGAGCACA CTTGCACTTC T 31

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GATCAGGAAG ATGAACGTAT TGTTCTGGTT GACAACAAGT GCAAGTCCGC TCGTATTACT 60

T

61

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 61 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTAGAAGTAA TACGAGCGGA CTTGCACTTG TTGTCAACCA GAACAATACG TTCATCTTCC 60

T

61

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 61 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GATCAGGAAG ATGAACGTAT TGTTCCTGGTT GACAACAAGT GCAAGGTTGC TCGTATTACT 60

T

61

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 61 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CTAGAAGTAA TACGAGCAAC CTTGCACTTG TTGTCAACCA GAACAATACG TTCATCTTCC 60

T

61

B66070"2FE30060

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 47 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CTAGAATCAT CCGTAGCTCA GAGGACCCAA ATGAAGATAT AGTCGAA

47

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 58 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GATACGGATG TTACGTTCTGA CTATATCTTC ATTTGGGTCC TCTGAGCTAC GGATGATT

58

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 49 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGTAACATCC GTATCATCGT CCCACTGAAT AACCGGGAGA ATATCTCAG

49

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 49 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

000533 010990
005070 0150050

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

49

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 49 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

49

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

49

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

44

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

33

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 60 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

60

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

44

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

44

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

60

62.

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

60

62

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GATCTGTGTA AGAAGTCTGA TATCGATGAA GATTCCGCTA CAGAAACCTG CAGCACATG 59

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

AATTCATGTG CTGCAGGTTT CTGTAGCGGA ATCTTCATCG ATATCAGACT TCTTACACA 59

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

GATCTGTCTA AGAAGTCTGA TATCGATGAA GATTACAGAT TCTTCAGACT ATAGCTACTT 60

CTAA 64

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

0005310-01099
B66070-9FE50060

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

AATCTTCATC GATATCAGAC TTCTTAGACA

30

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GATCTGGTTA AGAAGTCTGA TATCGATGAA GATTACCAAT TCTTCAGACT ATAGCTACTT

60

CTAA

64

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

AATCTTCATC GATATCAGAC TTCTTAACCA

30

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTGTCCAGC TCTACCTCTG TTGGATCACA CTTCTTACAC A

41

09005318-010999

(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 46 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

ACTCAAAGCA ACATTTGCGA TGAGGACAGC GCTACAGAAA CCTGCA

46

(2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 57 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GGTTTCTGTA GCGCTCTGCT CATCGCAAAT GTTGCTTTGA GTCGCAGTGA CTATCTG

57

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 59 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GCACCTACGA TAGGAACAAA TGCTACACGG CCGTGTTCC GCTCGTGTAT GGTGGAGAG

59

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

09005319 010998
866040 8450060

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GAGCGGAACC ACGGCCGTGT AGCATTTGTT CCTATCGTAG GTGCTGCA

48

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

ACAAAAATGG TGGAAACTGC CCTTACGCCC GATGCATGCT ATCCGGACTG

50

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

AATTCAGTCC GGATAGCATG CATCGGGCGT AAGGGCAGTT TCCACCATT TGTCTCTCC

60

ACCATACAC

69

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

0505318 040998

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

CTAGAAGTAA TACGAGCACA CTTGCACTTG GAGTCAGCGA CGGCCTTCTC TTCGGCGAAC 60

AGCCTCGGGT CTTGGATGGC AGCGACCT 88

(2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TGGTACGAAT TCCAGGTSMA RCTGCAGSAG TCRG 34

(2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

ACAGATATCG GGATTTCTCG CAGACTC 27

(2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

ACAGAATATC GTCAACACCT TCCCACCC 28

(2) INFORMATION FOR SEQ ID NO:87:

0005318 010990

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACAAAGCTTT TATTTACCCG ACAGACGGTC

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GTCCCCCCTC GAGCGAYATY SWGMTSACCC ARTCT

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACACTGCAGC AGTTGGTGCA GCATCAGC

28

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CTGCAGGAAG CGGAAGCGGA GGAAGCGGAA GCGGAGGAAG CGGAAGCGAA TTC

53

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

CCTTCGCCTT CGCCTCCTTC GCCTTCGCCT CCTTCGCCTT CGCTTAA

47

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

ACAGGATCCA TGGAAACCCC AGCGCAGCTT CTCTTCCTCC TGCTACTCTG GCTCCCAAGA

60

TACCACCGGA CCCGGG

76

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

355010"BT250060

TGGTACAGAT CTAGGTSMAR CTGCAGSAGT CRG

33

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

ACAGGAATTC AATTTTCTTG TCCACCTT

28

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

GTTCTAGAGA YATYSWGMTS ACCCARTCT

29

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

ACACCGCGGC AGTTGGTGCA GCATCAGC

28

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 75 base pairs
 - (B) TYPE: nucleic acid

05005318-010998

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

(2) INFORMATION FOR SEQ ID NO:98:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

(2) INFORMATION FOR SEQ ID NO:99:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 77 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GCGATGACGA CGATAAGGCC CAAACGGAGA CCTGTACTGT TGCGCCTCGT GAACGGCAAA 60
ACTGCGGATT CCCGGAA 77

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 66 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GTTTTGCCGT TCACGAGGCG CAACAGTACA GGTCTCCGTT TGGGCCTTAT CGTCGTCATC 60
GCTTCA 66

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 72 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GTAACACCCT CTCAGTGC GC TAATAAAGGC TGCTGTTTTG ATGACACGGT ACGGGGCGTT	60
CCGTGGTGCT TC	72

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 72 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GCCCCGTACC GTGTCATCAA AACAGCAGCC TTTATTAGCG CACTGAGAGG GTGTTACTTC	60
CGGGAATCCG CA	72

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 49 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

TACCCCAATA CAATTGACGT TCCGCCTGAA GAAGAGTGCG AGCCGTAAG 49

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 68 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

AATTCTTACG GCTCGCACTC TTCTTCAGGC GGCAAGTCAA TTGTATTGGG GTAGAAGCAC 60
CACGGAAC 68

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Pro Leu Gly Ile Ile Gly Gly
1 5

(2) INFORMATION FOR SEQ ID NO:110:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

Ile Ile Gly Gly
1

(2) INFORMATION FOR SEQ ID NO:111:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Val Arg Asp Gln Ala Gln Glu Asn Arg Ala Ser Gly Asp Ala Gly
1 5 10 15

Ser Ala Asp Gly Gln Ser Arg Ser Ser Ser Ser Lys Val Leu Phe
16 20 25 30

(2) INFORMATION FOR SEQ ID NO:112:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

0005348 040998

Val	Pro	Ser	Thr	Pro	Pro	Thr	Pro	Ser	Pro	Ser	Thr	Pro	Pro	Thr
1				5					10					15
Pro	Ser	Pro	Ser	Cys	Cys	His	Pro	Arg	Leu					
16				20					25					

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

Glu Gln Lys Leu Ile Ser Glu Asp Leu
1 5